

**COMPTE-RENDU
SEANCE NUMERO 1
INTRODUCTION :
ENJEUX DE L'IA EN
SANTÉ
FEVRIER 2023**

Séance 1 : Introduction

Intervenants :

- **Gérard Raymond**, président de France Assos Santé
- **Lydie Canipel**, co-présidente de la Société française de santé digitale (SFSD)
- **Jean-Jacques Zambrowski**, délégué général de la Société française de santé digitale (SFSD)
- **Emmanuel Bacry**, directeur scientifique du Health Data Hub, Directeur de Recherche au CNRS à l'Université Paris-Dauphine
- **Dominique Cardon**, professeur de sociologie à Sciences Po, directeur du Médialab

1. Gérard Raymond, président de France Assos Santé

Présentation de France Assos Santé

- La loi du 4 mars 2002 a été le point de départ de la reconnaissance des droits des patients en établissant la participation des associations agréées d'usagers.
- L'association France Assos Santé est l'union nationale des associations agréées du système de santé. Sa création par la loi Touraine de juillet 2016 constitue une reconnaissance de ces associations d'usagers de la santé qui œuvrent depuis 20 ans pour promouvoir la parole, la participation et la co-construction avec les usagers de la santé.

Il est indispensable d'impliquer les citoyens et usagers du système de santé dans les évolutions numériques du secteur. Les responsables du département du numérique au ministère de la santé ont saisi ces enjeux dès 2017 en associant France Assos Santé à ces réflexions.

Pour réussir cette transition vers le numérique et l'intelligence artificielle, la **confiance** est indispensable, et elle doit être bâtie en s'appuyant sur 3 piliers : la **participation**, la **transparence** et la **sécurité**.

La **participation** implique que chaque usager de la santé renforce ses compétences et s'approprie les informations concernant sa santé pour prendre, avec les professionnels de santé, les décisions adaptées pour lui-même. Il y a aujourd'hui un manque d'informations vulgarisées auprès des citoyens pour comprendre ces nouveaux dispositifs. Ce manque d'explications peut provoquer de la méfiance. La **transparence** est donc fondamentale pour instaurer la confiance.

Les données personnelles de santé permettent au **patient** d'améliorer le traitement de sa maladie par l'équipe soignante et en font un véritable **acteur de son parcours médical**. Il est donc indispensable de susciter l'adhésion pour obtenir une participation active des usagers.

En plus d'une meilleure prise en charge individuelle, **les données font progresser la recherche tout entière** et rendent le système de santé plus efficient dans son ensemble. On pense à la recherche sur des maladies, à une meilleure organisation du système de santé, ou encore au développement de nouveaux outils numériques utilisant ces données.

L'échange entre associations, décideurs, et concepteurs doit se poursuivre dans l'élaboration de ces dispositifs. L'annonce de la création d'un comité usagers-citoyens qui ouvre le champ de la participation est à ce titre une très bonne nouvelle.

Des moyens devront être mis à disposition pour améliorer l'appropriation de ces outils par les patients. La **fracture numérique** devra être résorbée : 7 millions de personnes sont encore privées d'un accès satisfaisant au numérique.

La **sécurité des données** de santé est un défi majeur pour bâtir un système fondé sur nos valeurs de solidarité et d'humanisme. C'est d'abord à l'État et à la représentation nationale d'édicter les lois et règlements prévoyant des sanctions fortes et dissuasives vis-à-vis d'acteurs ne respectant pas ces valeurs. Le rôle des associations d'usagers est ici déterminant pour questionner et améliorer ces textes qui nous protègent.

France Assos Santé croit fermement dans le numérique en santé, tout en restant très attentif à sa mise en œuvre pratique. Le secteur de la santé a beaucoup à gagner dans cette transformation.

Questions :

Comment expliquer l'évolution des associations d'usagers sur la question des outils numériques et plus particulièrement du partage de données de santé ?

Cela s'explique par la reconnaissance et la représentation des associations de santé dans la construction de ces dispositifs. On assiste à une évolution de la démocratie en santé et on ne peut que se féliciter de la volonté des pouvoirs publics d'impliquer les usagers.

Y a-t-il une spécificité de l'IA dans ces enjeux liés au numérique ? Est-on face à une prolongation de ces problématiques ou une rupture ?

Il s'agit du même processus : ces logiciels doivent toujours être construits en associant les usagers. Les données et leur sécurisation restent des questions centrales. L'IA vient simplement concrétiser ces enjeux de bonne utilisation des données.

2. Lydie Canipel, co-présidente de la Société française de santé digitale (SFSD), et Jean-Jacques Zambrowski, délégué général de la SFSD

Présentation de la SFSD

- La SFSD est une société savante en santé digitale, constituée de groupes d'experts

- Créée en 2008, la SFSD a pour ambition de constituer un espace de réflexion sur l'évolution des pratiques professionnelles médicales impactées par le numérique.

En 2008, la SFSD était très avant-gardiste. À l'époque, il y avait une méfiance face à la numérisation de la santé, en particulier une crainte de déshumaniser la médecine et de couper le patient de son médecin. La santé digitale n'était pas une spécialité médicale et elle ne l'est toujours pas aujourd'hui. Un chemin progressif a été fait depuis, impacté par des textes de loi et un contexte réglementaire en forte évolution.

La SFSD porte la conviction que le numérique est au service d'une **santé humaniste** : cette valeur est constitutive de la pratique médicale. Le rôle de la SFSD est d'accompagner les professionnels de santé et les citoyens dans l'évolution des pratiques professionnelles pour continuer à exercer ces métiers de façon humaniste.

Dans la couverture médiatique du numérique en santé, ce sont surtout les **outils** qui sont mis en avant. On oublie trop souvent l'**impact sur les pratiques professionnelles** et le travail d'adaptation que devront fournir professionnels de santé et aux patients.

Les professionnels de santé sont trop souvent négligés dans ces transformations et ont une pratique du numérique par défaut. Le combat de la SFSD est de leur donner les moyens d'exercer leurs métiers avec des outils adaptés.

Plusieurs évolutions en cours montrent l'actualité de ces réflexions. L'arrivée du décret télésoin⁶ marque une sortie du tout présentiel et va donner un sens au parcours coordonné qui va devenir un parcours alterné. Le professionnel choisit la façon la plus efficace pour soigner son patient, en alternant prise en charge en présentiel et en distanciel par des équipes pluriprofessionnelles. L'espace numérique de santé amène les professionnels de santé à réfléchir à la bonne préparation des citoyens sur ces sujets. La SFSD se place dans une démarche d'accompagnement de ces évolutions.

Questions :

Quels sont les changements apportés par l'intelligence artificielle dans les pratiques des professionnels de santé ?

La SFSD a voulu se doter d'un département IA. Il ne faudrait pas recommencer les mêmes erreurs de non-préparation des professionnels de santé et des patients dans l'appropriation de l'intelligence artificielle. Il est important de dissiper la peur d'une substitution des médecins par des outils intelligents.

La diabétologie est un bon exemple : on dispose aujourd'hui de logiciels experts utilisant des algorithmes. Ces logiciels intelligents déclenchent des alertes, et souvent le professionnel de santé

⁶ Le décret télésoin de juin 2021 fait entrer dans le droit commun le télésoin pour 18 nouveaux professionnels de santé. Voir <https://solidarites-sante.gouv.fr/soins-et-maladies/prises-en-charge-specialisees/telesante-pour-l-acces-de-tous-a-des-soins-a-distance/article/le-telesoin>

ne fait pas confiance à l'algorithme qui a généré l'alerte. Il refait l'analyse des courbes glycémiques déjà effectuée par l'outil avant de poser son diagnostic. Nous voudrions accompagner les professionnels de santé pour créer de la confiance dans ces algorithmes : l'outil doit prendre en charge une partie de son travail pour le laisser se consacrer à son rôle d'expert et lui faire gagner du temps. L'apprentissage de cette confiance dans les algorithmes doit passer par une co-construction d'outils testés et validés avec les professionnels de santé.

Il faut maintenir un contrôle humain. L'algorithme peut négliger ce qui est aux marges pour ne retenir que ce qui est le centre de la courbe de Gauss : c'est l'une des limites du système, l'outil ne peut pas remplacer le médecin.

Qu'est-ce que la crise sanitaire a changé par rapport aux enjeux du numérique en santé ?

Avec la crise sanitaire, nous sommes passés d'une utilisation confidentielle à un recours massif à la téléconsultation qui a satisfait un tiers des médecins et une large partie des malades. Cette adoption s'est faite dans un contexte de nécessité : des modes dégradés de consultation ont été acceptés à titre dérogatoire. Dans la majorité des cas, cela n'a pas entraîné une dégradation de la prise en charge en elle-même.

La formation des médecins est importante : la SFSD a mis en place des formations exigeantes pour les professionnels de santé et mis à disposition des contenus de sensibilisation pour les usagers afin de découvrir cette pratique de la médecine.

La crise du covid a apporté un regard neuf de la tutelle ministérielle : la puissance publique veut encourager les techniques numériques en médecine.

Quels sont les besoins de formation des professionnels sur ces sujets ?

Il y a bien des besoins de formation, mais les professionnels ne se rendent que très peu aux formations proposées. On observe en particulier des freins pour les spécialités médicales : en cardiologie, la part de téléconsultation est de 0,2%, contre 5,9 chez les généralistes. Il y a bien un manque d'accompagnement à l'évolution des pratiques professionnelles et un problème d'appariement entre besoin et demande de formation.

3. Emmanuel Bacry, directeur scientifique du Health Data Hub, Directeur de Recherche au CNRS à l'Université Paris-Dauphine

Les dates jalons du Health Data Hub

- 29 mars 2018 : à la suite de la remise du rapport Villani, Emmanuel Macron annonce la création d'un hub des données de santé
- Juillet 2019 : l'article 41 de la loi « Ma Santé 2022 » prévoit la création du Health Data Hub
- 1^{er} décembre 2019 : création du Health Data Hub en remplacement de l'Institut National des Données de Santé, qui en reprend et élargit les missions

La France dispose de **bases de données** de santé très riches et nombreuses, mais elles restent largement **sous-utilisées**, pour 3 raisons principales :

- Les données sont dispersées
- Leur gouvernance et les procédures d'accès sont lourdes et hétérogènes
- L'Interopérabilité entre les bases de données reste très faible

Le Health Data Hub veut rendre possible l'exploitation de données de santé pour la recherche d'intérêt public. Le Hub est un **groupement d'intérêt public**, avec 59 parties prenantes réparties en 9 collèges : État, Caisse d'Assurance Maladie, Organismes d'assurance maladie complémentaires, recherche et enseignement, Établissements de santé, professionnels de santé, usagers, agences et autorités publiques indépendantes, industriels.

Le Hub sera un guichet unique pour l'accès aux données de santé, une plateforme sécurisée et scalable, un catalogue de données construit de manière progressive, et favorisera la mise en relation des acteurs clés du secteur.

En 2016 est créé le **Système national de santé** (SNDS) sous la responsabilité de la CNAM. Cette évolution majeure a officialisé la réunion de 3 bases :

- le SNIIRAM : base de la carte vitale contenant les données de remboursement
- PMSI : base traçant le parcours hospitalier
- CépiDC : base de causes médicales de décès

La loi santé 2019 élargit ce système national des données de santé. Ce SNDS élargi couvre tous les soins qui sont au moins partiellement remboursés par l'Assurance maladie, ce qui recoupe 90% des soins en France. Cette base est donc beaucoup plus complète et moins biaisée que les bases publiques de pays comme les États-Unis, où la protection sociale publique couvre moins de soins.

Le Hub n'est pas une base de données unique : il s'agit d'un fonctionnement en **catalogue de bases de données**. Les données vont être déposées sur la plateforme du Hub, mises en qualité et chaînées pour relier les bases. L'accès aux données s'effectuera depuis un espace sécurisé et ne concernera que les données explicitement demandées et pseudonymisées. Des frameworks standards d'IA seront intégrés dans la plateforme pour l'exploitation des données.

L'accès est gratuit pour la recherche académique et payant pour les acteurs privés. La sécurisation est très stricte, avec des audits fréquents. Le Hub permettra des gains substantiels au niveau des possibilités de croisement des sources et de la mise à disposition de capacités élastiques de calcul et de stockage. La première version du catalogue sera publiée à l'automne 2021.

200 chercheurs sont aujourd'hui affiliés, et le Hub accompagne 47 projets pilotes. Quand le catalogue sera publié, des projets pourront être déposés au guichet du hub pour obtenir l'accès à des bases du catalogue. Ces premiers projets pilotes permettent de tester le fonctionnement du Hub et accompagnés par la plateforme. Voici quelques exemples :

- La startup Implicity travaille sur le suivi en temps réel de pace maker pour prédire des complications cardiaques.
- Des projets d'analyse d'images avec des outils de deep-learning :

- Un projet multicentrique sur des IRM de prostate qui tirera parti de la puissance de calcul du Hub
- Un projet d'analyse des textes libres de comptes rendus médicaux

La liste des projets est disponible sur le site du Hub. Tous les projets ne sont pas de l'intelligence artificielle : il y a beaucoup de choses à faire sans IA, notamment à partir de chaînage de bases et de statistiques descriptives.

Les accès aux données du catalogue sont strictement encadrés : le CESREES évalue la qualité scientifique du projet, puis une autorisation de la CNIL est sollicitée. Les études réalisées seront partagées sur le site du Hub.

Questions :

Quels sont les champs les plus actifs aujourd'hui dans le développement de l'IA en santé ?

La reconnaissance d'image est actuellement particulièrement dynamique. À moyen terme, l'IA pourra jouer un rôle dans le croisement de données hétérogènes. Par exemple mélanger de l'imagerie avec de la génomique et des comptes rendus médicaux.

Dans les choses qui sont actuellement plus difficiles, l'IA a encore du mal à obtenir des résultats pour ce qui touche à l'optimisation du parcours de soin. Pourtant ce sont des problèmes et des types d'algorithmes communs à de nombreux secteurs hors de la médecine, par exemple dans la vente (modélisation d'un parcours client).

Le Hub va-t-il intégrer l'espace numérique de santé ?

Le Hub échange étroitement avec l'ENS pour essayer de construire une interopérabilité, mais il y a des difficultés à surmonter. Le Hub doit nécessairement fonctionner avec des pseudonymes, alors que l'ENS est par définition directement identifiant, il faudra donc un acteur entre le Hub et l'ENS.

Pour les producteurs de données dont le modèle de financement est de monétiser l'accès aux données, comment va s'organiser le partage de la valeur ?

C'est un sujet extrêmement important et sensible. Pour le moment les choses ne sont pas entièrement actées. Il est acquis que l'accès sera gratuit pour les académiques et payant pour les industriels, les tarifications sont encore en cours de construction.

Où sont hébergées les données ?

En Hollande, et Microsoft joue le rôle d'hébergeur. Il n'y aura pas les outils Microsoft dans l'espace sécurisé. Les données sont cryptées et décryptées dans chacun des espaces sécurisés et ce cryptage est géré indépendamment de Microsoft.

Une étude comparative a été menée pour aboutir à cette décision. Pour aller vite, la seule solution réaliste était Azur. Les autres solutions impliquaient d'importants coûts et temps de développement

pour obtenir des garanties de sécurité suffisantes. Le gouvernement est arrivé à cette décision avec l'accord de l'ANSI.

Le Health Data Hub a été conçu dans une optique de réversibilité : dès qu'une option « souveraine » existera, la bascule sera faite sur cette nouvelle plateforme.

Le risque de sécurité ne sera de toute façon jamais absolument nul et il doit être mis en balance avec les risques liés à la non-ouverture des données.

L'intégration dans une plateforme européenne commune est-elle à l'agenda ?

Cela est très compliqué, car les juridictions sont différentes dans chacun des pays.

4. Dominique Cardon, professeur de sociologie à Sciences Po, directeur du Médialab

Qu'entend-on par « IA » ? Au début des années 1990 la recherche avait enterré l'IA : le mot était vu comme mal choisi, ces technologies ne marchaient pas, beaucoup d'argent avait été investi sans résultats probants. Les années 2012-2013 marquent un tournant avec l'essor de quelque chose qui ne s'appelait pas encore de l'IA à l'époque : les **réseaux de neurones**. La situation d'aujourd'hui est aussi un **effet de mode** : on habille sous le terme « IA » toute une série de choses, alors que du point de vue de l'histoire de cette famille de technologies ces choix paraissent incohérents.

L'histoire de l'IA est souvent représentée comme une alternance entre des phases de fort intérêt pour ces technologies (financements publics, médiatisation, ...) suivies de phases de désintérêt (les « hivers de l'IA »). Ces 3 vagues se seraient produites dans les années 1950-1960, puis dans les années 1980, et enfin dans les années 2010. Cette représentation est une mauvaise manière de raconter l'histoire des sciences car elle donne un aspect évolutionniste et continu à ce qui est en réalité une **rupture**.

Le clivage fondamental dans cette histoire réside dans l'opposition, encore vivace aujourd'hui, entre un courant **symbolique** et **connexionniste**, deux projets différents pour rendre les machines intelligentes. Le courant **symbolique** est historiquement associé aux deux premières vagues de l'IA. Il cherche à programmer la machine pour qu'elle produise des raisonnements en fonction de symboles et de façon logique. Ce courant n'a pas fait de percées absolument décisives. Le projet **connexionniste** s'appuie sur l'idée de faire un calcul distribué inspiré des neurosciences. Ce principe qui remonte aux années 1940 a été mis à la porte avec l'arrivée de « l'IA » dans les années 1950 et l'avènement du symbolisme qui souhaitait scripter la fonction cognitive.

Une autre manière de raconter l'histoire de l'IA est de montrer un **changement de paradigme entre le connexionnisme et le courant symbolique**. Ce qu'on a vu réapparaître en 2010, lors de la « troisième vague », n'a rien à voir avec l'IA symbolique des deux premières vagues et revient aux traditions de la cybernétique en fonctionnant sous la forme d'un réseau de neurones. Cette technologie est aujourd'hui appelée IA, mais elle ne l'était pas du tout jusqu'à très récemment.

On a beaucoup réanimé l'IA avec le robot, la fiction de la machine autonome alors que c'est exactement le contraire qui s'est passé en réalité : dans les représentations communes et médiatiques de l'IA, on a invoqué un imaginaire symbolique alors que globalement ce sont les technologies connexionnistes qui décollent.

Ce courant connexionniste ne raisonne pas et ne calcule pas de symboles, il définit l'intelligence comme perceptive. Contrairement aux symbolistes, le projet connexionniste ne programme pas à l'avance : **c'est à l'algorithme lui-même de trouver ses propres règles**, et plus les données en entrée sont fines, meilleur sera le modèle.

Il a été question de co-construction aujourd'hui, et ce point est délicat avec ces technologies. Le courant connexionnisme fait le pari qu'il faut connecter l'IA avec des données **sans modèle a priori** : ce n'est pas le concepteur qui va le fabriquer.

Quand la base de données n'est pas d'une grande richesse, un réseau de neurones ne sert pas à grand-chose. Emmanuel Bacry l'a bien dit : avec les données du Health Data Hub, on peut mettre en œuvre beaucoup d'applications autres que des réseaux de neurones. Quand les données ne sont pas assez fines, l'IA (au sens de deep learning) n'est pas forcément la solution la plus adaptée.

Un autre enjeu central est celui de connecter le modèle perceptif des réseaux de neurones avec des approches plus traditionnelles. Se pose donc la question d'articuler les bases de données classiques avec les bases sur lesquelles l'IA travaille.

Les débats sur l'explicabilité sont très actifs, et il faut garder en tête que la transparence d'un algorithme ne donne pas une explication de son fonctionnement dans le cas des réseaux de neurones : les concepteurs de ces outils ne les comprennent pas, même s'il est possible de retracer une partie de leur fonctionnement.

Questions :

Quels sont les obstacles à l'explicabilité ?

Le problème est que les réseaux de neurones sont non linéaires. L'explication mathématique est encore à trouver, et l'explication de la manière dont ces données ont produit telles prédictions est rendue impossible dès lors qu'on a des centaines de milliers de poids (les vecteurs d'un modèle sont les poids des relations entre les couches de neurones).

Les réflexions sur l'éthique de l'IA doivent donc se concentrer sur d'autres points : bien décrire les données, identifier très clairement l'objectif de l'outil, créer des systèmes ad hoc pour vérifier sur des variables de contrôle qu'il n'y a pas de biais.

Comment traiter la question de la personnalisation de ces outils à la grande diversité de situation des patients ?

C'est une question centrale : comment avoir un raisonnement statistique qui maintient la multi-dimensionnalité des données ? Un système classique de régression statistique va se concentrer sur le parcours de soin moyen. La prétention des approches connexionnistes est qu'on peut emprunter

différents chemins pour aller vers le même objectif. On peut donc maintenir une personnalisation : cette promesse est au cœur de ces dispositifs.

Dans le numérique, c'est bien ce qui est appliqué, on peut penser aux systèmes de recommandation sur les plateformes musicales. Il y a bien l'idée que le système est adaptatif : le système se sert des comportements pour réviser la manière dont est produit un modèle. On est face à une boucle d'apprentissage qui peut être personnalisée pour l'utilisateur.

Doit-on faire le deuil de la transparence ? Peut-on trouver des substituts ou d'autres critères de contrôle de ces algorithmes ?

Cela dépend. On dit « l'IA », mais en réalité il y a une grande diversité d'algorithmes qui n'ont pas les mêmes caractéristiques.

Ensuite, la transparence n'est pas la même chose que l'explicabilité. Rendre les algorithmes transparents est une préoccupation toujours d'actualité. La question est d'arriver à expliquer ce qu'ils font. On ne peut plus avoir une explicabilité complètement interne à ces systèmes, mais on peut tester des modèles en faisant de la rétro-ingénierie : cela permet de re-simuler le fonctionnement des calculs. On vérifie ainsi que des biais ne sont pas présents. Il ne faut pas abandonner le problème politique des exigences d'explication des algorithmes. Il faut aussi garder à l'esprit que les problèmes d'explicabilité de notre environnement socio-technique ne sont pas l'apanage de l'intelligence artificielle.